DOCUMENT RESUME

ED 121 119 FL 007 637

AUTHOR Just, Marcel Adam; Carpenter, Patricia A.
TITLE Eye Fixations and Cognitive Processes.
INSTITUTION Carnegie-Mellon Univ., Pittsburgh, Pa. Dept. of Psychology.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.; National Inst. of Mental Health (DHEW), Bethesda, Md.
PUB DATE Aug 75
GRANT NIE-G-74-0016
NOTE 71p.; For related documents, see FL 007 636, FL 007 632 and EJ 112 015

EDRS PRICE MF-$0.83 HC-$3.50 Plus Postage
DESCRIPTORS *Cognitive Measurement; *Cognitive Processes; *Eye Fixations; Eye Movements; Eyes; Memory; Neurolinguistics; Psycholinguistics; *Psychological Studies; Reading Processes; Task Performance; Verbal Learning; Visual Perception

ABSTRACT

This paper reports on a study concerned with rapid mental operations of the central processor as it performs tasks such as the comparison of rotated figures, mental arithmetic, sentence verification, and memory scanning. The central processor is the site of most of the symbol manipulation that takes place in the human information processing system. One of the goals of the study is to demonstrate that the locus, duration and sequence of eye fixations are all closely tied to the activity of the central processor. The research is intended to discover how eye fixations are related to cognitive processes, the aim being to construct a theory of that relationship that generalizes across a number of task environments. The second goal of the research is to examine how the eye fixation data reveals the fine structure of the processor's activity in performing a number of cognitive tasks. (Author/CLK)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED 022 619

DEPARTMENT

OF

PSYCHOLOGY

Carnegie-Mellon University

EYE FIXATIONS AND COGNITIVE PROCESSES[1]

Marcel Adam Just

and

Patricia A. Carpenter

Psychology Department

Carnegie-Mellon University

Pittsburgh, Pennsylvania 15213

Complex Information Processing

(CIP) Paper #296

August, 1975

Abstract

This paper presents a theoretical account of the sequence and duration of eye fixation during a number of simple cognitive tasks, such as mental rotation, sentence verification, and quantitative comparison. In each case, the eye fixation behavior is linked to a processing model for the task by assuming that the eye fixates the referent of the symbol being operated on.

A widely accepted view of the human information processing system is that most of the symbol manipulation takes place in a central processor, sometimes referred to as the active memory (Neisser, 1967), working memory (Newell & Simon, 1972), operational memory (Posner, 1967), or the executive processor (Newell, 1973). This paper is concerned with the major mental operations of the central processor, and how they are reflected by the patterns and durations of eye fixations during a task involving visual input. We will examine the basic operations, parameters, and overall structure of the central processor as it performs such tasks as the comparison of stored figures (Shepard & Metzler, 1971), mental arithmetic (Parkman, 1971), sentence verification (Carpenter & Just, 1975), and memory scanning (Sternberg, 1969). These tasks generally take less than 5 or 10 sec to complete, and can be decomposed into many rapid mental operations, often estimated to execute between 10 to 500 msec each. The goals of this paper are (1) to demonstrate that the locus, duration, and sequence of eye fixations are closely tied to the activity of the central processor, and (2) to examine how the eye fixation data reveals the fine structure of the processor's activity in performing a number of cognitive tasks.

Our proposal is that the eye fixations the referent of the symbol currently being processed if the referent is in view. That is, the fixation may reflect what is on the "top of the stack." If a number of symbols are processed in a particular sequence, then their referents should be fixated in the same sequence, and the duration of fixations on each referent may indicate how long the corresponding symbol is operated on. The obvious advantage of monitoring eye fixations is that the behavior within any particular trial can potentially be decomposed into various stages where hypotheses can be directly measured. By contrast, a single response usually cannot be interpreted as decomposed within reference to processes in other conditions.

[illegible] [illegible] [illegible] [

Results. The mean response latencies for correct Same trials, shown in Figure 2, increased monotonically with increasing angular disparity. All three subjects showed a linear increase between 0° and 100°, but the curves were positively accelerated beyond 100°. The mean latencies from 0° to 100° disparity have a pattern similar to that obtained by Metzler and Shepard (1974).

Insert Figure 2 about here

Eye fixation results. One striking feature of the eye fixation behavior was that subjects systematically looked back and forth between the left and right figure.[2] For example, at 0° disparity, subjects initially fixated the left figure, then looked over at the right hand figure, then looked back at the left, and frequently looked back at the right hand figure for a second time, for a total of three switches between the two figures. The mean number of such switches between figures increased with angular disparity, as shown in Figure 3.

Insert Figure 3 about here

The next step of the analysis was designed to determine exactly what subjects were looking at, and how the pattern of their fixations might reveal the microstructure of the underlying cognitive operations. To classify the locus of the eye fixation, we divided each figure into three main segments: the arm whose third face of the end cube was visible (open), the arm whose third face of the end cube was not visible (closed), and a central angle. For example, in Figure 1a, the upper arm will be called the open arm, while the lower arm will be called the closed, and the four central cubes constitute the central angle. The locus of the eye spot was scored according to the locus of its centroid with respect to the three segments.

The simplest way to describe our scoring procedure is to apply it to a

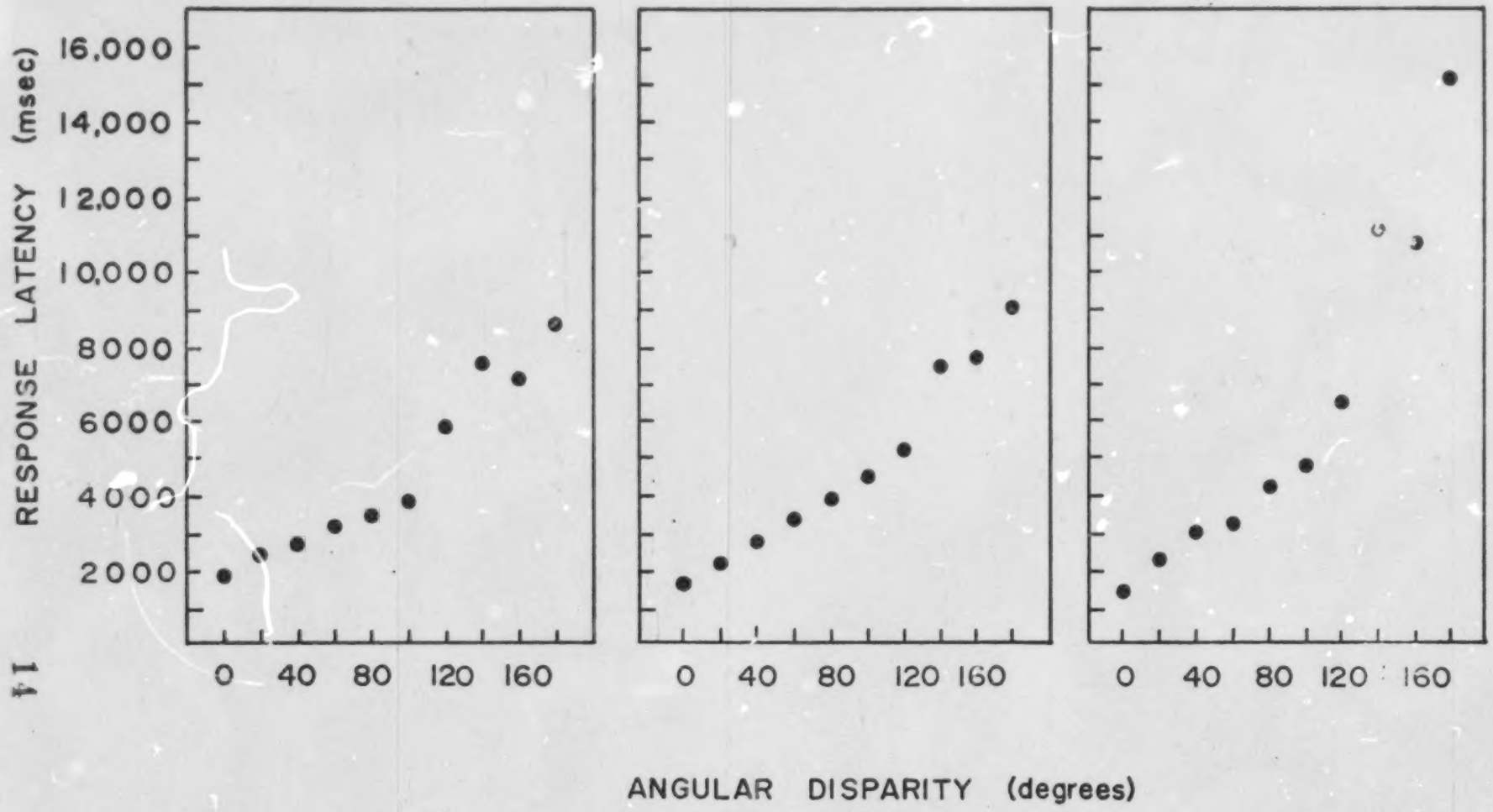


Figure 2. Mean response latency for Same trials as a function of angular disparity for the 3 subjects.

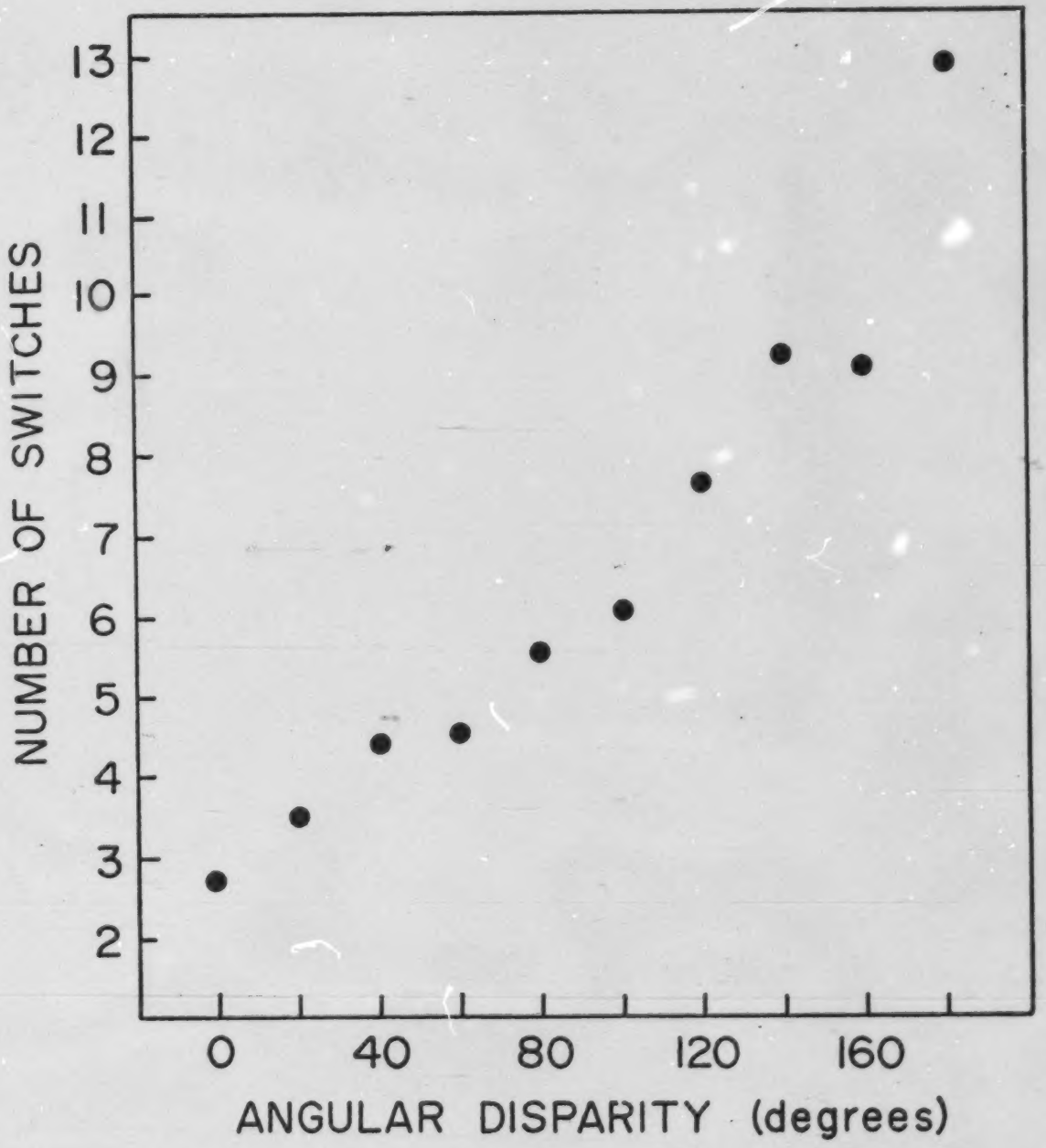


Figure 3. Mean number of switches for Same trials as a function of angular disparity.

few representative scan paths. Figure 4 shows a scan path for a Same trial with 0° disparity. After the initial fixation on the center of the left figure, the subject fixated corresponding closed arms at the upper part of each figure. Then the open arms at the bottom of each figure were fixated.

Insert Figure 4 about here

To make our analysis of the scan paths precise, we constructed rules for classifying instances of search, transformation and comparison, and confirmation. The most prominent property of the scan paths was that the subject would repeatedly look back and forth between corresponding segments of the two figures without any intervening fixations on other segments. We identified the repeated fixation of corresponding segments with the transformation and comparison process. When the same pair of segments was involved in two transformation episodes separated by other fixations, their durations were combined. The transformation and comparison process is evident in fixations 5 to 8 of the scan path shown in Figure 5, where the figures have an 80° disparity. In fixations 5 to 8, the subject looked back and forth between the closed arms of the two figures, for a total of 1185 msec.

Insert Figure 5 about here

We identified the search process with the initial portion of the scan path that preceded the first instance of transformation. Applying these rules to the scan path in Figure 5, fixations 1 to 4 would be attributed to search, for a total of 818 msec. In Figure 4, where the angular disparity is much smaller, the duration of the search process (351 msec, fixation 1) is much shorter.

We identified the third process, confirmation, as a short sequence of fixations between corresponding parts of the two figures other than the trans-

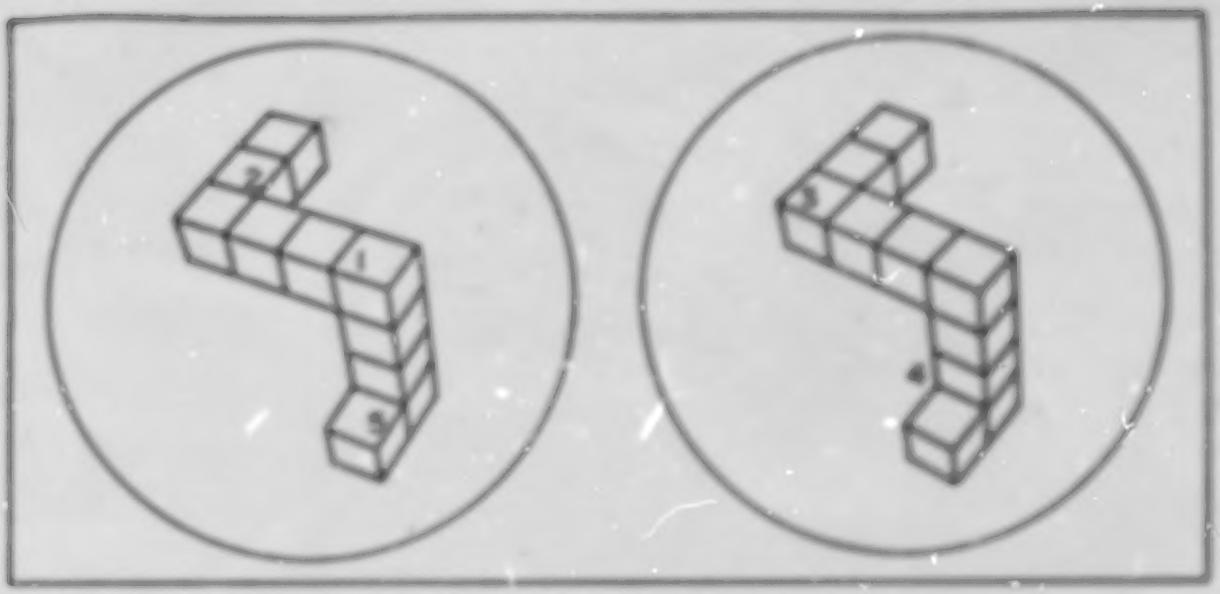


Figure 4. The figure indicates the sequence of fixations on a correct Same trial in which the disparity was 0°. The subject's total response latency was 1296 msec, of which 11% had no visible eye spot. The locus and duration of the fixations are as follows:

Fixation	Figure	Location	Duration
1.	Left	Center	351 msec
2.		Closed arm	150 msec
3.	Right	Closed arm	200 msec
4.		Open arm	200 msec
5.	Left	Open arm	250 msec

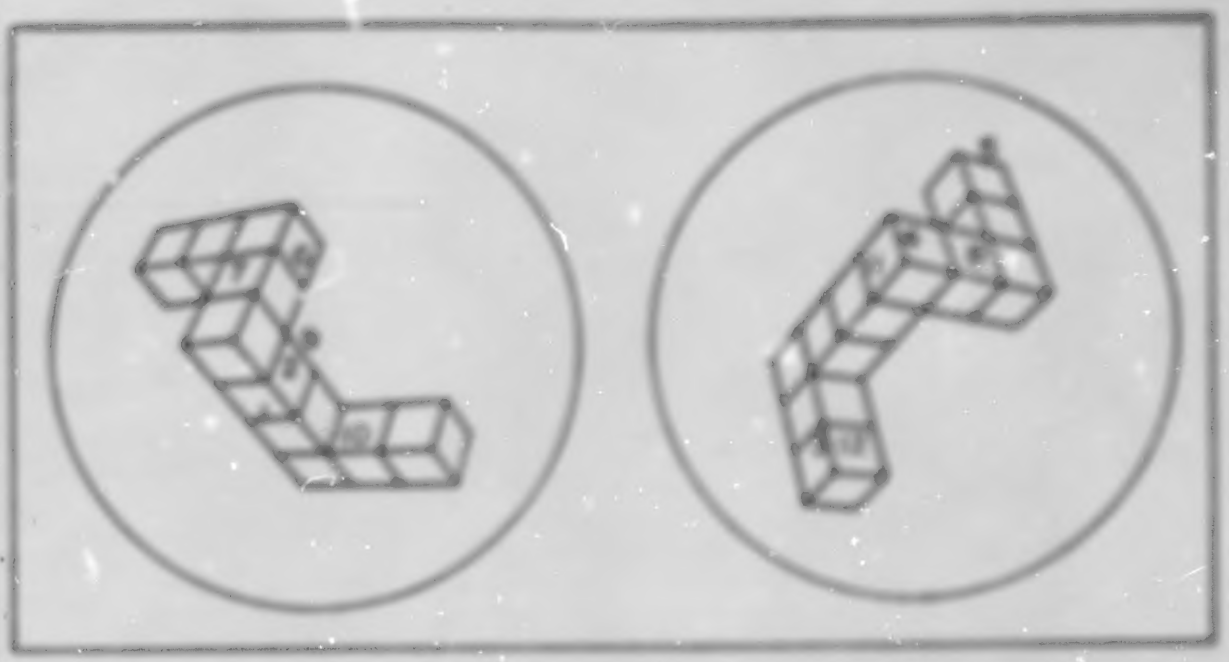


Figure 1. The figure indicates the sequence of fixations on a correct Same trial in which the disparity was 80'. The subject's total response latency was 3174 msec, of which 8% had no visible eye spot. The locus and duration of the fixations are as follows:

Fixation	Figure	Location	Duration
1.	Left	Center	200 msec
2.		Center	301 msec
3.	Right	Open arm	167 msec
4.		Center	133 msec
5.		Closed arm	167 msec
6.	Left	Closed arm	200 msec
7.		Closed arm	317 msec
8.	Right	Closed arm	301 msec
9.	Left	Center	133 msec
10.		Open arm	350 msec
11.	Right	Center	400 msec
12.		Open arm	317 msec

[illegible] segments. Confirmation could appear on a scan from the central eight to six o'clock on one figure, then a similar scan on the other figure. Figure 8 shows an example of the confirmation process where the fixations proceed from the central eight to the eyes and on the left figure (fixations 9 and 10) and then a similar scan is executed on the right (fixations 11 and 12). In the scan path in Figure 8, the last two fixations on the eyes area (fixations 9 and 5) also exemplify confirmation.

To see how well the model fits the eye fixation data, we scored for each pattern of 100 of the (7) correct base trials. Seventy-one trials were not scored because of apparatus failure, in which the optical system failed to capture an eye spot that was visible at least 80% of the time. The scan patterns [illegible] from the 100 total sample are very similar to the data for all (7) scored 4 base trials, as shown in Figure 9, so the sample appears to be representative.

Insert Figure 9 about here

The analysis of the eye gaze allows us to examine how the total processing time is distributed across search, redistribution and comparison, and confirmation stages as a function of angular disparity. As Figure 9a shows, the time spent initially searching the figures increased with angular disparity, from about 300 msec at 0° to about 1600 msec at 180°. The bulk of the processing time was spent in recognition and comparison, as shown in Figure 9b. The duration of this stage increased markedly with increasing angular disparity, from about 600 msec at 0° to 2500 msec at 180°. The minimum time spent in the third stage, confirmation, increased from 400 msec at 0° to 2500 msec at 180°, as shown in Figure 9c. Thus, for a typical trial, say at 60° disparity, 2/5 of the time was devoted to initial search, 3/5 to recognition and comparison, and 2/5 to confirmation. The remaining 140

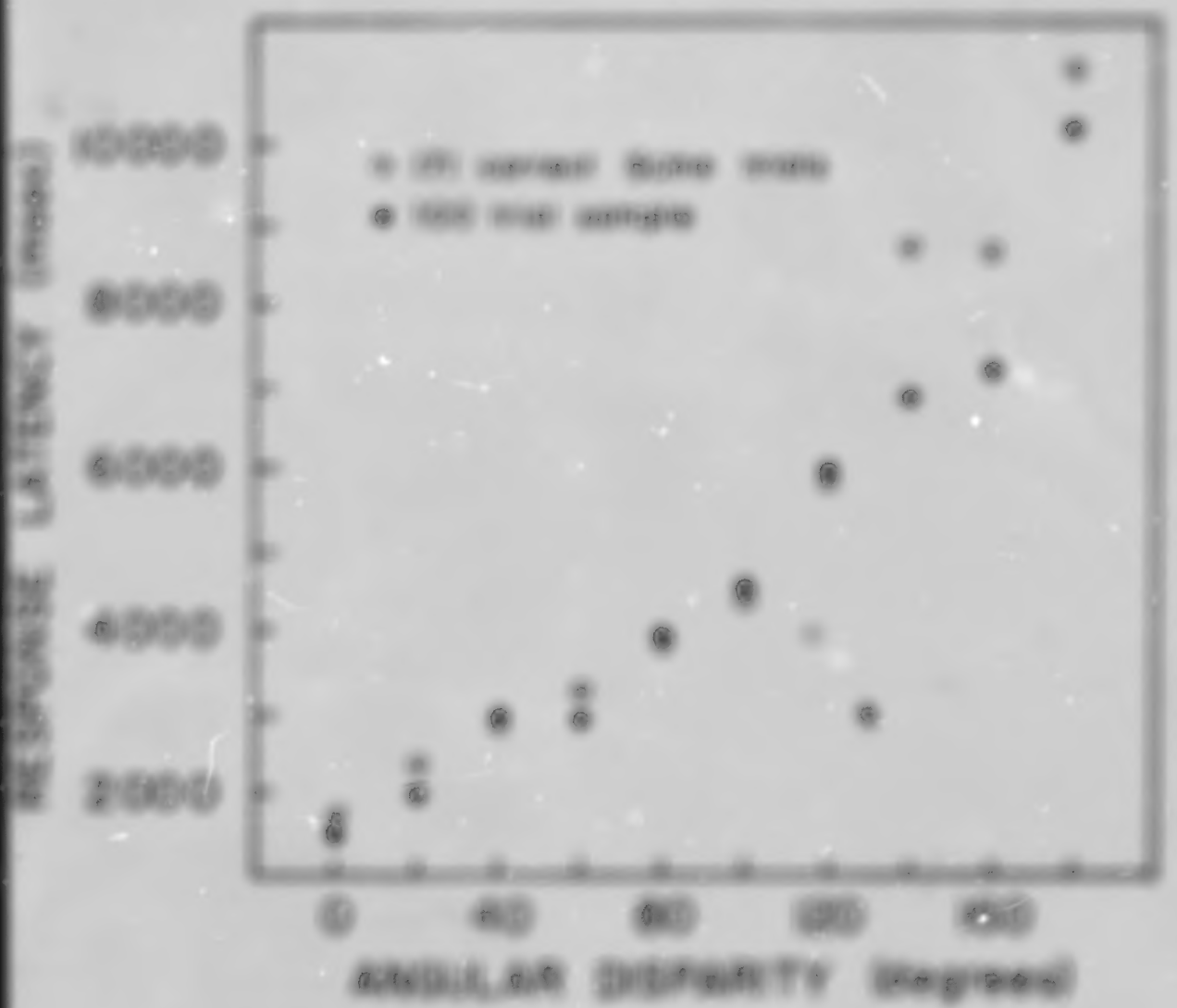


Figure 6. [illegible]

[illegible]

[illegible]

[illegible]

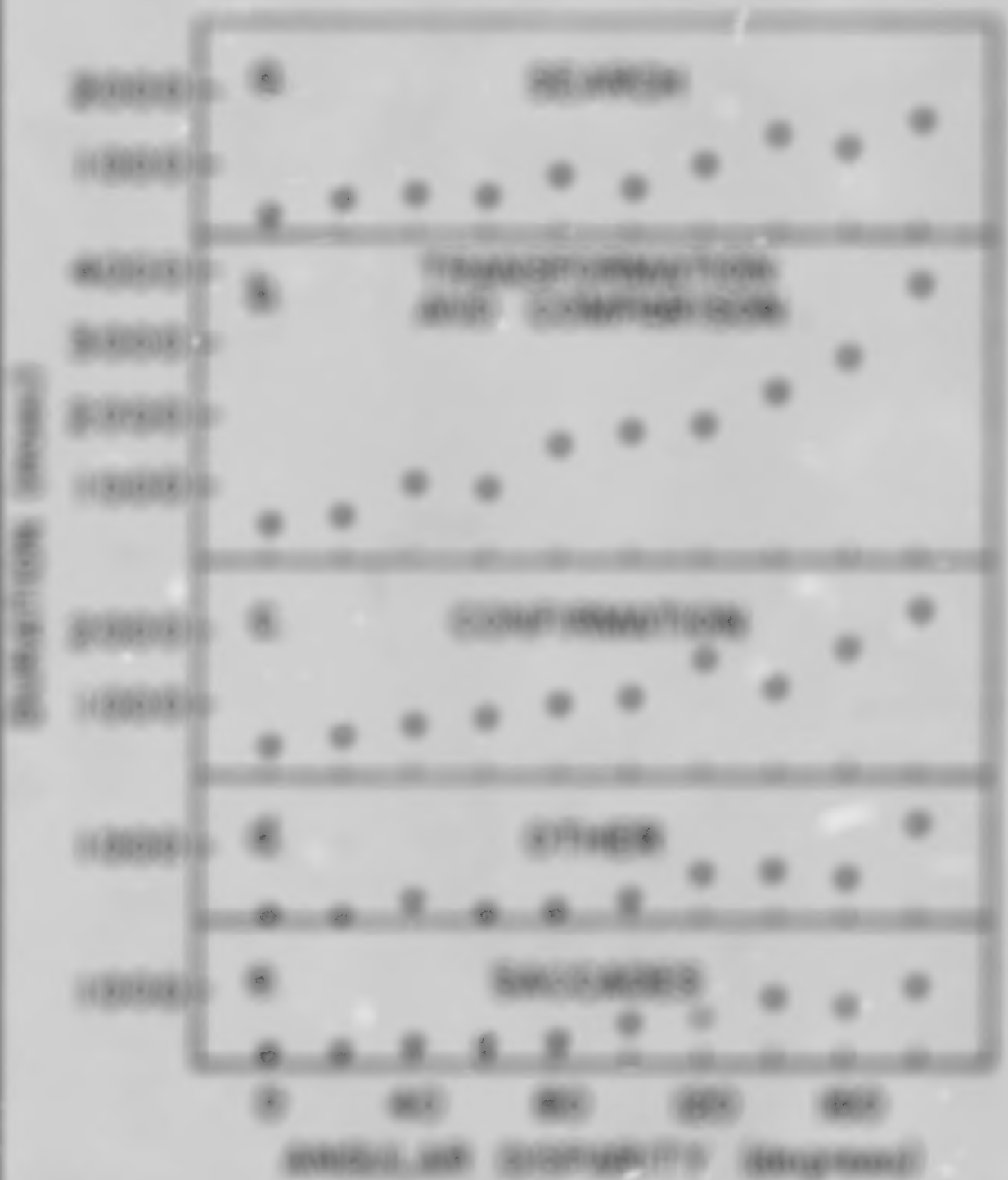


Figure 7 [illegible]

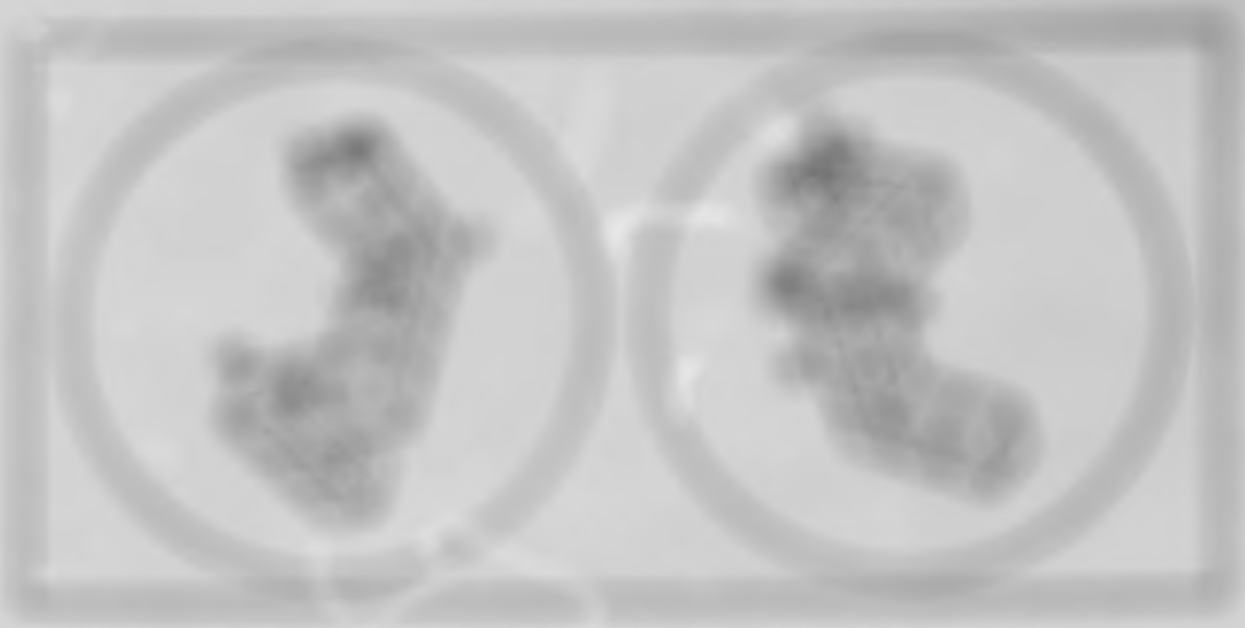
[illegible]

[illegible]

[illegible]

[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

[illegible]



[illegible]

since it leads to the discovery that the inter-segment relations are not the same in the two figures and hence that the figures are different. In fact, one of the most prominent features of the Different scan paths is the large amount of confirmation behavior that they contain. In the ten analyzed Different trials, the confirmation process consumed an average of 4195 msec, or 49% of the total duration.

The prolongation of the confirmation process is not the only reason for the very long response latencies for Different trials. On some trials, after going through a complete search-transform-and-unsuccessfully-confirm sequence, subjects make a second attempt at searching, transforming and confirming a different pair of segments. Occasionally, a lengthy search stage involved an examination of all the possible ways of pairing the segments, and that kind of search led directly to a response of Different, without any transformation. Thus, the durations of all three stages increased during Different trials, but the duration of confirmation increased the most.

One scan path that exemplifies the processing on Different trials is shown in Figure 9. Fixations 1 to 4 reflect the initial search for corresponding segments, consuming 1436 msec. Then, there is a transformation and comparison of the short arms of each figure in fixations 5, 6, and 7, consuming 919 msec. Fixations 8 to 15 reflect the confirmation process, consuming 3175 msec. We presume that it is during confirmation that the subject determined that the relation between the arm and central angles was different in the two figures. In this trial, the bulk of the processing time was consumed by the confirmation stage.

Insert Figure 9 about here

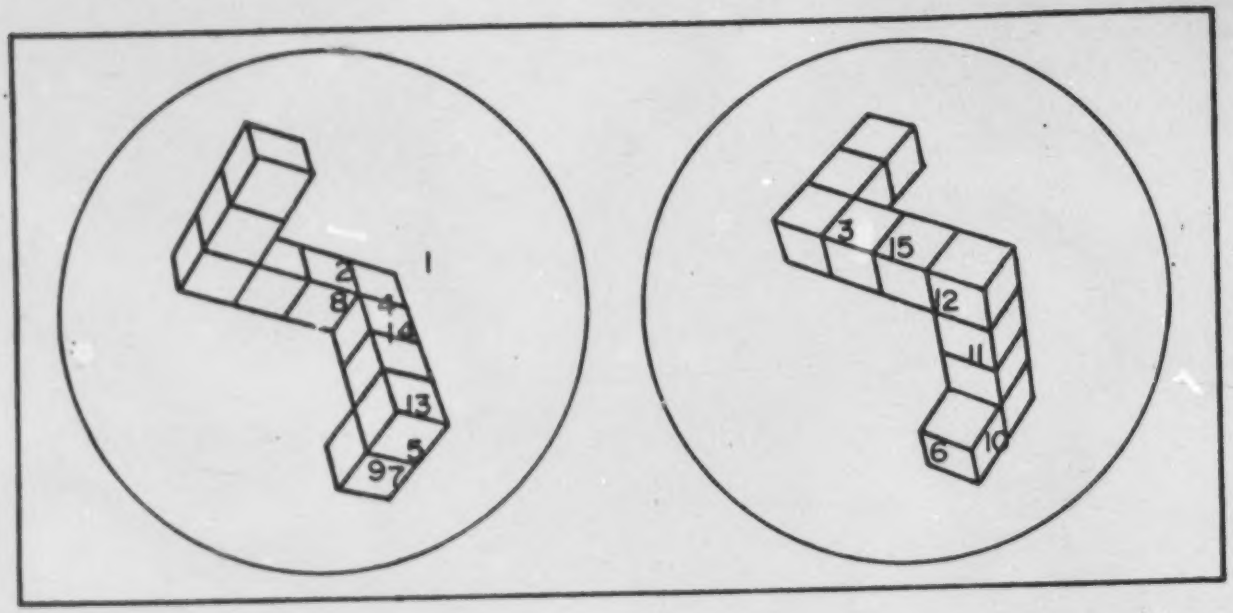


Figure 9. The figure indicates the sequence of fixations on a correct Different trial. The subject's total response latency was 5868 msec of which 6% had no visible eye spot. The locus and duration of the fixations are as follows:

Fixation	Figure	Location	Duration
1.	Left	Center	234 msec
2.		Center	518 msec
3.	Right	Closed arm	367 msec
4.	Left	Center	317 msec
5.		Closed arm	184 msec
6.	Right	Open arm	434 msec
7.	Left	Closed arm	301 msec
8.		Center	251 msec
9.		Closed arm	217 msec
10.	Right	Open arm	635 msec
11.		Center	518 msec
12.		Center	234 msec
13.	Left	Closed arm	585 msec
14.		Center	167 msec
15.	Right	Center	568 msec

The Processing Model of the Rotation Task

The internal representation. We propose that the processor operates on one segment of the figure at a time, and that the representation of the segment is schematic. The representation must include information about the segment's absolute orientation in space, as well as some defining feature such as its length or whether it is a closed or open arm. This information can be efficiently represented as the vector formed by the major axis of the segment. Moreover, if the vector has its initial point at the origin of the reference frame, then the segment can be represented by the spherical coordinates of the end point of the vector. For example, an open arm might be represented (OPEN (r, θ, ϕ)) where r is the length of the segment, and θ and ϕ define the orientation of the segment.

The initial search process. The scan paths indicate that the search for corresponding segments uses a simple heuristic. Once a segment of one figure has been identified, then the search for the corresponding segment starts in the corresponding location of the other field. For example, if the long arm is in the upper-right hand corner of the left field, then the search for the corresponding segment begins in the upper-right hand of the right field. If there is no segment in the upper right, then the segment nearest the upper right is examined. The duration of this search process increases with angular disparity for two reasons. First, with increasing disparity, the corresponding segments are in successively more dissimilar locations. At 0° disparity, corresponding segments have identical locations in their respective fields. However, as the disparity increases from 0°, absolute location is a successively poorer cue for finding corresponding segments, and the heuristic must be supplemented by an active search. The second reason for the increase is that at larger disparities, the prob-

ability of selecting and attempting to transform non-corresponding segments increases and this incorrect transformation is counted as part of initial search. Figure 8 shows an example of the search process selecting non-corresponding segments that are both at the top of their respective fields. In this trial, the incorrect search led to an error. On other trials, the incorrect selection of a pair of segments was detected after some transformation had been attempted. Thus, the eye fixations allow us to trace the initial search for corresponding features and to determine the reason for the increase in the duration of the search process with angular disparity.

The transformation and comparison process. The eye fixation data also allow us to formulate a precise model of the transformation process. We propose that rotations are executed and monitored in discrete steps of approximately 50°. The estimate of the 50° step size is based on the result that there is one additional switch during the transformation stage for each additional increment of 50° in angular disparity, as shown in Figure 10.[3] A transformation may consist of applying a rotation rule that alters the representation of the orientation of a segment by 50°. For example, an open arm represented as (OPEN ($\underline{r}$, θ, ϕ)) might be transformed into (OPEN ($\underline{r}$, $\theta + 50°$, ϕ)). It is assumed that the representations of the two segments are rotated towards each other by applying the 50° rotation rules first to one segment and then to the other, until they are within 25° of each other. This form of representation and transformation does not impose any great computational burden, in contrast to a truly analogue, holistic representation of the entire figure rotated by a parallel computation of the position of all its points.

Insert Figure 10 about here

This model of the transformation stage is most easily explained by working through an example, say when two Same figures have an angular disparity

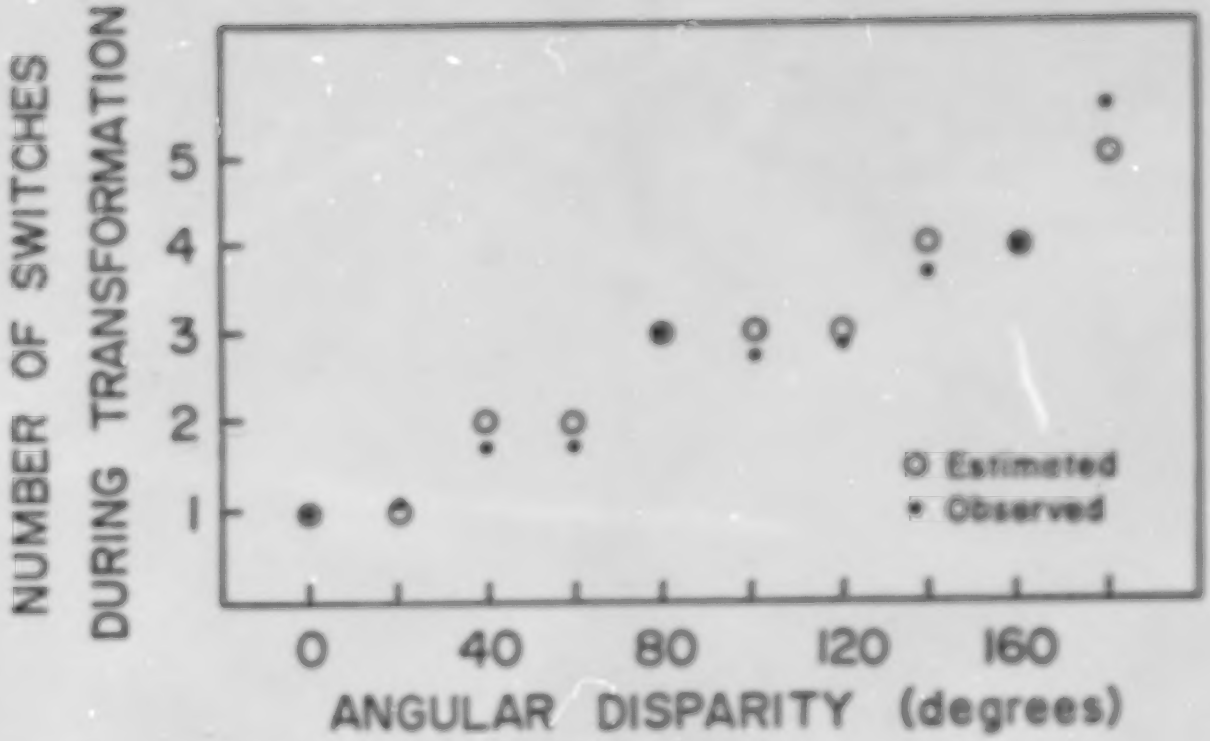


Figure 10. Mean number of observed and estimated switches during the transformation and comparison stage in Same trials as a function of angular disparity.

of 80°. Suppose that the subject has encoded a particular segment of the left-hand figure. The first switch occurs when she fixates and encodes the corresponding segment of the right figure. Then, the orientations of the two segments are compared. The orientations differ by more than 25°; therefore, she rotates her representation of the right-hand figure by 90° counterclockwise. After this transformation, she retrieves the representation of the segment on the left. In doing this retrieval, she switches her fixation to that figure. After it is retrieved, she compares the two orientations. They would still be more than 25° apart. Therefore, she transforms the representation of the orientation of the left-hand figure by 90° clockwise. Then she retrieves the representation of the segment on the right in order to compare the two again. In doing this retrieval, she switches fixation over to the right figure. After it is retrieved, she compares the two orientations. At last, after three switches and two applications of the rotation rule, the two segments are represented at fairly similar orientations (within 25° of each other). The subject would then continue on to the confirmation process.

This model can be easily summarized. After the subject has initially encoded one segment on a figure, she proceeds through the following steps:

1. Retrieve the representation of the corresponding segment on the other figure (and switch fixation to other figure).
2. Compare the two orientations. Are they less than 25° apart?
   a. No. Transform the currently fixated segment by 90° in the direction of the other figure. Go back to Step 1.
   b. Yes. Go on to the confirmation stage.

The model assumes a very close relationship between eye fixation and mental operations during the transformation process. The saccadic rule is always applied to the eye that is being fixated. Applying a saccadic rule to the representation of one eye causes the representation of the other eye to be pushed down in the short-term memory stack. When the representation is being retrieved to the top of the stack, the other eye is fixated next (Step 3). According to this model, the number of switches during transformation should increase monotonically with the angular disparity, but the increase should be in the form of a particular step function. There should be one switch between 0° and 25°, two switches between 30° and 70°, three switches between 75° and 120°, and so on. Figure 16 shows that this predicted number of switches corresponds very closely to the observed number of switches. The increase in switches is similar to the pattern obtained for the duration of the transformation process, shown in Figure 15. The data in Figure 15 suggest that each step of the transform and compare process takes about 800 msec. Thus, the model of the transformation stage gives a good account of the data.

The confirmation process. Detecting two segments into similar orientations during the transformation stage is not sufficient for deciding whether the figures are the same or different. Therefore, the third stage, confirmation, determines whether segments other than the transformed pair correspond to each other. The eye paths indicated two methods for confirming such correspondence. One method applies the same sequence of saccadic rules used in the transformation stage to another pair of segments. If this second rotation is successful, then the two figures are the same. This method, used on about half the trials, was identified by switches of fixation between a pair of segments other than the transformed pair. A second method verifies the relation between the central angle and an arm of each figure and deter-

given stimulus that evokes it) as well as both figures. Neurons such as the other half of the circle, are described to a unit that has cells to the use of an figure and that a similar one of the other figure (see fig- ure 2 for an example). The combination of these two periods would produce a confounded scenario that increases and explain disparity, but the increase would be at a slower rate than for the transformation neurons (shown in fig- ure 1). Other period of confinement can determine the response of cells to different.

Conclusion. We can therefore now look to a detailed model of the pro- cessing in the Shepard-Metzler task, but there are questions about the gener- alizability of the model. Without assessing a broad range of experimental situations, there is no way of knowing which aspects of the model are neces- sary of the ACorn processing current and which aspects are task-specific. Consider the proposed M^3 processes origin. It is possible that the M^3 origin are independent and associated over tasks. Alternatively, people may be able to vary the size of the invariance step to the particular pairs and stage of processing differences that are being made to be examined. This is a more empirical question of whether the invariance operation adapts itself to the task environment. Similarly, we can question whether the representa- tion of the figures is the same in all invariance tasks. The representations proposed by the current model are highly viewpoint, but may be unable with fixated information to perform the task. The representations ought to even respond to tasks that require that more information to succeed than the Fig- ures. But as the Fixation succeeds but in a positive tasks for the Shepard- Metzler task, this methodology should also distinguish the invariant from the invariant processes, and so lead to a general theory of mental rotation.

Mental rotation has been investigated as an analogue process (Kosslyn & Shepard, 1974) in the sense that creating a representation that can retain

[illegible] to another magazine from the implementation goes through web browser process that corresponds to conventional environments. Although Bennett and Bergeron [1975], [illegible] pen pointing to convert from performance years differences can be modified to [illegible] the 45° page that is prepared in related image. It is possible that entire such 90° page based per transmission into per corresponding to [illegible] [illegible]. But one and 90° angle, a 180° processor becomes [illegible] magazine corresponding to 90° and 180° in network, and as the proposed process is to a [illegible] image.

In addition, the new public method is to approach the processing that search, transformation, and conference images and to decrease the [illegible] of each image. Because to Kramer design the transformation image before half that the process was increased to image of approximately 90°. This [illegible] not applicable are easy to the correct [illegible] image, but also get [illegible] problems as prior image and different images. The research shows that our Kramer can [illegible] the sequence of device generation during the processing of special information.

[illegible] Software and Systems

Under the parameters that crowd give software processing device per network, we can study the project search [illegible] a process to take to [illegible] of an [illegible] process. In particular, as our [illegible] has the most processing time to [illegible] among the various images of verification. Many time-time studies of wireless verification time that pages web and [illegible] and time longer to improve when switching a negative [illegible]. The wide processing time for a negative time between 100 and 1000 ms, depending on the [illegible] processes of the negative wireless [illegible] to [illegible]. While the processing image [illegible] a verification [illegible] reading the wireless will increase [illegible], in using the process, and capacity for the [illegible] of the process to the [illegible] from the wireless [illegible].

[illegible] [illegible] [illegible] [

finer components. For this analysis, we divided the gazes into four categories: the initial gaze on the sentence, subsequent gazes on the sentence after having looked away, gazes on the location specified by the directional term in the sentence, and finally, gazes in any other locations. Thus, the durations of all four types of gazes add up to the total response time. The important question was whether these durations varied systematically as a function of the four information conditions.

The initial gaze on the sentence should reflect the time to read and represent the sentence. As Figure 13 shows, the duration of the initial fixation was 57 msec longer for negatives than for affirmatives, F (1, 33) = 14.93, $p < .01$. This result indicates that the negative sentences take about 57 msec longer to read and represent than the affirmatives. After having looked away from the sentence, subjects occasionally refixated it later in the trial. The durations of such subsequent gazes on the sentence were similar for all four information conditions, as Figure 13 shows.

Insert Figure 13 about here

The directional term in the sentence can be viewed as an instruction for where to direct the next fixation--irrespective of whether the sentence was affirmative or negative. In fact, the location specified by the directional term was the locus of the second gaze on 92% of the trials. Subjects tended to fixate this location only once during a trial. The time spent gazing at this location increased linearly with the number of hypothesized comparison operations, as Figure 13 shows. The straight line accounts for 98.1% of the variance among the four means. The residual 1.9% is not significant, F (2, 33) <1. This suggests that after reading and representing the sentence, the subject fixated the picture and completed the comparison operations while fixating the location specified by the directional term. The slope, 135 msec per operation, may be interpreted as an estimate of the

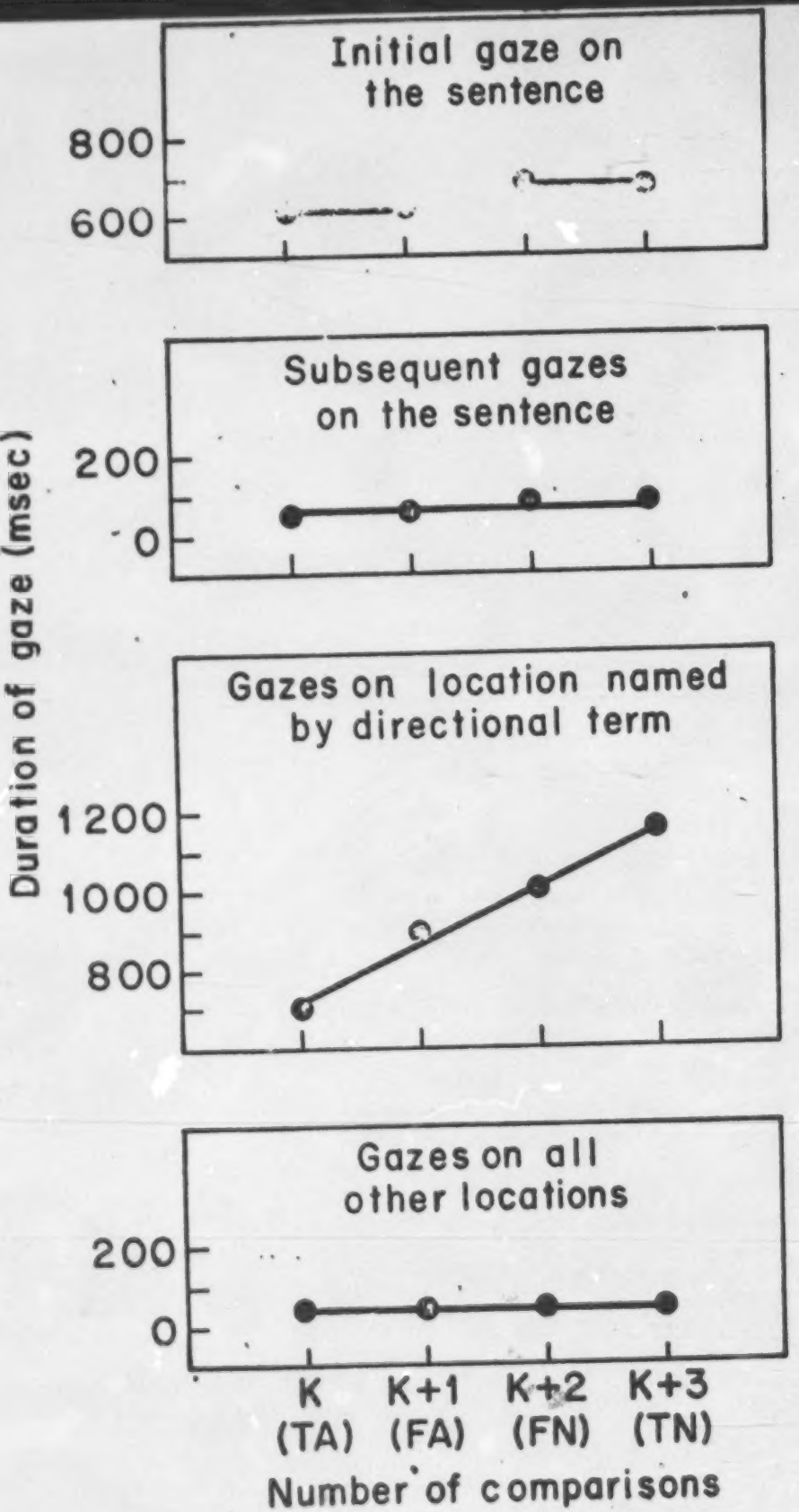


Figure 13. Average duration spent gazing on various locations of the display for the true affirmative (TA), false affirmative (FA), false negative (FN), and true negative (TN) conditions. These components add up the total response times shown in Fig. 12.

time to compare an element from the sentence representation to one from the picture representation.

Occasionally, subjects would gaze at a location other than the sentence or the location specified by the directional term. The frequency and duration of these other gazes did not vary as a function of information condition, as Figure 13 shows.

These results indicate how the total processing time in sentence verification is distributed among various stages. The duration of the initial gaze on the sentence suggests that the time needed to read and represent the sentence is 700 msec at most. This 700 msec enters primarily into the intercept of the total response time. What accounts for the difference between the response time for the fastest condition, the 1400 msec for the true affirmative, and the slowest condition, the 1900 msec for the true negative? This 500 msec is consumed by the operations that compare the sentence and picture to determine their relation. In fact, these comparison operations are reflected in the duration of the gaze on the location specified by the directional term.

This analysis can tell us why negative sentences take longer to process than affirmatives. The total response time was 346 msec longer for negatives than for affirmatives. This can be partitioned into several components. The largest component is the comparison time (reflected in the duration of gaze at the picture) which was 267 msec longer for negatives. Secondly, negative sentences took 57 msec longer to read. And thirdly, subsequent gazes on the sentence were an insignificant 20 msec longer for negative sentences. Thus the bulk of the additional processing time for negatives is consumed by the operations that compare the information from the sentence to the picture.

The results show that there is a systematic correspondence between the mental operations and eye fixations in a sentence verification task. Under well-controlled conditions, the sequence of gazes on the external display corresponds to the sequence of mental operations in the processor. Moreover, the duration of the gaze is proportional to the duration of the underlying operations.

QUANTITATIVE COMPARISON

A quantitative comparison requires an order judgment (e.g., Which is larger?, or Which is brighter?, or Which is longer?) of two or more objects along a common underlying dimension. The comparative judgment requires that the two objects be represented, and their representations be compared. In order to obtain more detailed evidence about the processes in this task, we devised an experiment in which subjects' eye fixations were monitored while they decided which of two groups of dots was larger. The response latencies for selecting the larger of two groups of dots strongly resemble the latencies for digit comparisons (Buckley & Gillman, 1974), so our task may produce results generalizeable to digit comparisons. Furthermore, prior data (summarized by Klahr, 1973) have shown that the time to determine how many dots there are in a group increases monotonically from about 500 msec for one dot, to 2200 msec for nine dots. These results suggest that larger groups of dots might be fixated longer if they are to be quantified. We hypothesized that the duration of fixation on each of the groups of dots might tell us how the two groups of dots were represented and processed during a quantitative comparison task.

Method. Subjects' eye fixations were monitored as they compared the sizes of two groups of dots. Each group contained from 1 to 6 dots, so there were 15 possible pairs of unequal groups. If the word more appeared on the

left side of the display (as shown in Figure 14), subjects indicated whether the upper or the lower group contained more dots, by pressing an upper or lower response button. If the word was less, they judged which group contained fewer dots. A total of 60 stimuli were formed by orthogonally combining the two words, more and less with the 15 pairs of groups and the responses designating either the upper or the lower group. Each subject had four blocks of 60 stimuli, presented in a random order. A trial started 500 msec after the subject fixated a point at the locus of the word, and pressed a "ready" button.

Insert Figure 14 about here

The computer-generated display was presented on a video monitor at a distance of 53 to 68 cm. The word more or less, 2.8 cm wide, appeared 13 cm to the left of the dot display. The dots formed two vertical lines one above the other, separated by a vertical distance of at least 3 cm. Each group of dots was .5 to 6 cm long, depending on the number of dots in the group. For scoring purposes, the screen was divided into the four imaginary sectors indicated by the dashed lines in Figure 14. The analysis was concerned primarily with the distribution of the gaze across the four sectors.

Results. The response latencies showed that this experiment replicated the major latency results that have been previously reported for this task (Buckley & Gillman, 1974). The mean latencies ranged from 700 to 1100 msec. We will discuss the response latencies in more detail after an analysis of the eye fixation results.

First of all, we examined the duration of gaze on the smaller group. If subjects were computing the number of dots in the group, one might expect that the more dots there were in the group, the longer people would spend looking at it. As expected, the gaze duration on the smaller group in-

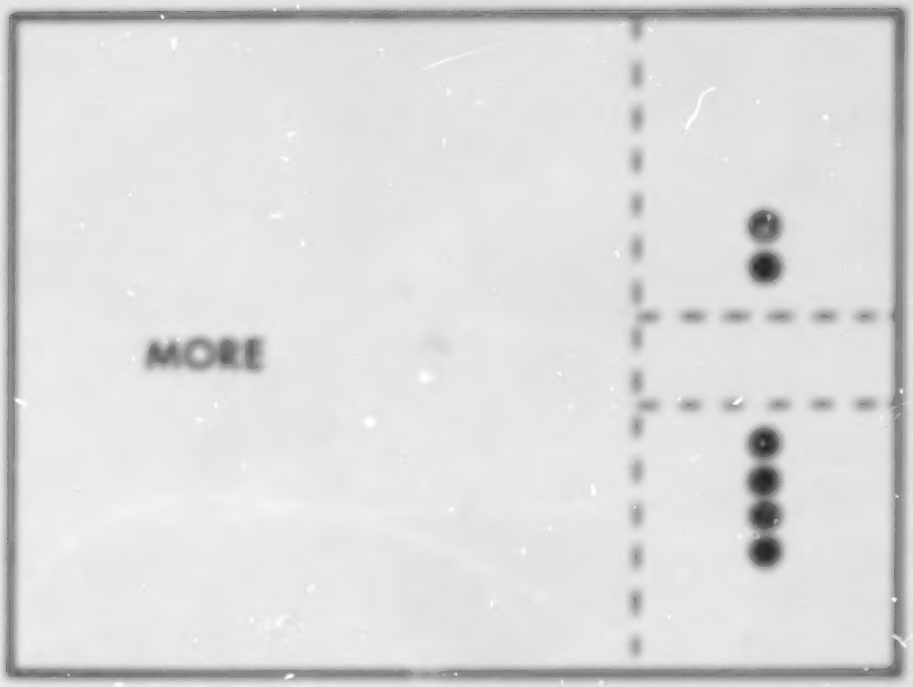


Figure 14. Schematic diagram of the visual display in the dot quantitative comparison task (not to scale). The dashed lines which indicate the boundaries between the four sectors, did not appear in the display.

crossed by about 30 msec for each additional dot (see Figure 15). This is within the range of subitizing rates cited by Klahr (1973), although it is at the low end. Thus, it is possible that the subjects counted the number of dots in the smaller group.

Insert Figure 15 about here

But what about fixation on the longer group? If subjects estimate the quantity of dots in the longer group, then gaze duration on the longer group should also increase with the number of dots in that group. However, Figure 14 shows that the duration of gaze on the longer group is independent of the number of dots there. Thus, the two groups of dots are fixated differently. The size of the smaller group predicts the gaze duration in the smaller group, but the size of the larger group does not predict the gaze duration on the larger group.

Insert Figure 16 about here

The proposed model. The results are consistent with a counting model (cf. Fuson, 1971; Groen & Parkman, 1972) adapted to the dot inequality task. The process might start by counting one or two dots in both groups, and checking to see if either group has been exhausted. If one group has been exhausted, it would be designated the smaller one. If neither has been exhausted, then one or two more dots might be counted in each group, and again there would be a check to see if either group has been exhausted. This process would continue until one of the groups, the smaller one, would be exhausted. If the subjects were answering the question "Which group contains more dots?", they could simply indicate the group that had not been exhausted. The number of counts or increments in this process would be proportional to the number of dots in the smaller group. If the gaze duration is proportional to the number of increments, then it follows that duration

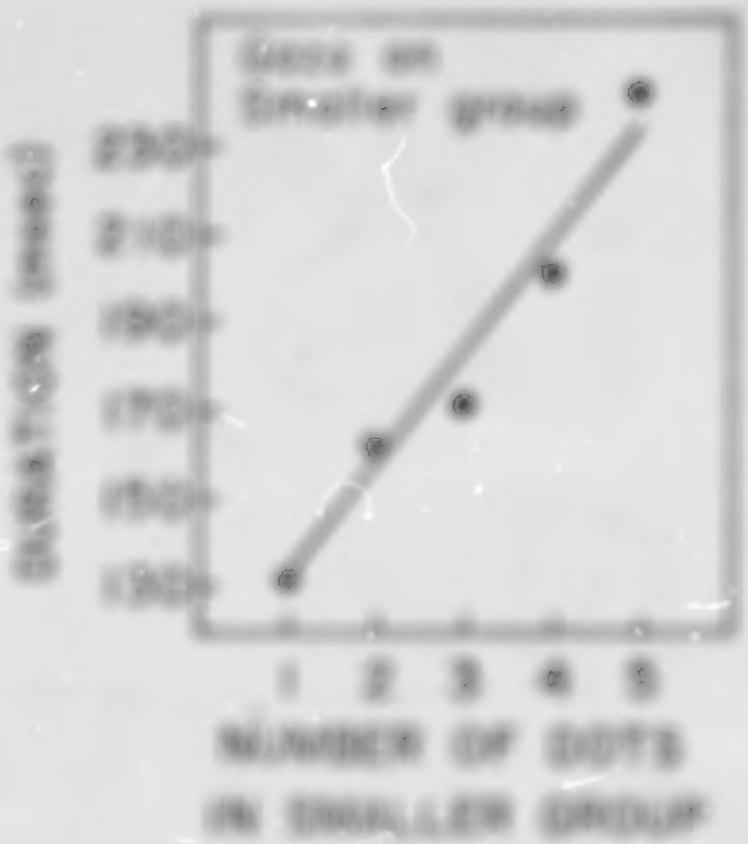


Figure 14. Mean duration of gaze on the emotion group of dots as a function of the number of dots in the [illegible].

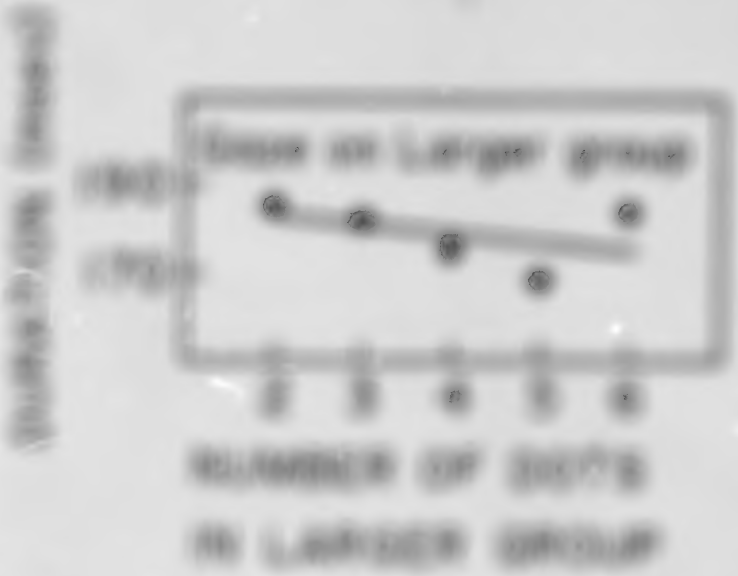

Loss in Larger group
(80)
(70)
2
3
4
5
6
NUMBER OF DOTS
IN LARGER GROUP

Figure 1x. [illegible]

[illegible]

followed by a picture, and examine how the internal representation of the prior sentence alters the way the picture is scanned in a verification task. For example, we have used this methodology to examine the processing of affirmative and implicitly negative sentences (Carpenter & Just, 1972). The affirmative sentences (e.g., A small proportion of the dots are red) and the implicitly negative sentences (e.g., Few of the dots are red) have the same truth value. However, linguistic and psychological evidence suggests that the two sentences have different internal representations (Just & Carpenter, 1971). The affirmative sentence is represented as an affirmation that the small subset has some property, in this case, redness. We predicted that after reading the affirmative that refers to the small subset (e.g., A small proportion...,, people should tend to fixate the small subset. By contrast, an implicit negative is represented as a negation of some property of the large subset, in this case, redness. We predicted that after reading an implicit negation about the large subset (e.g., Few of the dots are red), people would tend to fixate the large subset. As predicted, subjects looked at the location in the picture specified by the underlying representation of the sentence. The locus of the eye fixation is sensitive to the deep structure representation, even when subjects aren't consciously aware of the nature of the linguistic stimulus or of their pattern of eye fixation.

While people are listening to spoken questions or passages, they tend to fixate the pictorial referent of words that occur in the text (Cooper, 1974; Kahneman & Lass, 1971, cited by Kahneman, 1973). For example, in the Kahneman and Lass study, people

were shown a schematic drawing of four objects, such as a car, person, tree and airplane and asked a question like "What makes of cars can you name?" Subjects tended to look at the schematic car while answering. More interestingly, when the picture was removed prior to the question, subjects still tended to look where the appropriate object had been located. Such fixations apparently play a place-keeping organizational role rather than an encoding role. The symbols in the short term memory may be indexed to particular spatial locations. (This formulation is reminiscent of the method of loci (cf. Bower, 1970) and spatial interference effects in retrieval (Byrne, 1974)). When the time comes to retrieve or operate on a symbol, the eye may fixate the location from which the symbol was originally encoded. It may be this mechanism that produces fixations on the referent of the symbol at the top of the stack, assuming that the referent stays in the same location.

Duration of gaze. In the tasks we investigated, the time spent gazing at a figure reflected both the time to encode that figure as well as the time to operate on the encoded symbol. Tachistoscopic recognition studies indicate that familiar figures, like alphanumeric characters or even words can be internalized within a very short exposure duration--as low as a few tens of milliseconds. Yet in these cognitive tasks, people gaze at very simple and familiar figures for much longer, often for hundreds of milliseconds. For example, in the sentence verification task, subjects looked at a star (*) or a plus (+) for 700 to 1200 msec, depending upon the relation between the sentence and the figure. Clearly, the duration of the gaze includes not only encoding time but also the time for subsequent operation on the encoded symbol.

There are a number of reasons why a subject might continue to fixate a figure after the relevant information has been encoded. If the processor

is busy operating on the most recently encoded information, there is no reason for it to direct the eye to seek other information. So the eye may remain stationary simply because it is not instructed to move. An alternative view of the persistence of the gaze is that the processor might actively instruct the eye not to move during the processing of the most recently encoded information. The reason for avoiding new fixations might be that a saccade automatically initiates an encoding activity (cf. Loftus, in press) that could interrupt the ongoing processing. Perhaps the reason that people often gaze upwards or close their eyes altogether while computing the answer to a demanding question is that they are avoiding extraneous encoding operations that could interrupt processing. Thus, the persistence of the gaze could be due to the absence of an instruction to move the eye or the presence of an instruction not to move the eye. In either case, the gaze duration on a particular figure provides a measure of the time spent processing the corresponding symbol.

One of the most elegant studies of the relationship between gaze duration and mental operation examined gaze duration in a Sternberg memory-scanning task (Gould, 1973). In Gould's experiment, subjects had a memory set of 1, 2, or 3 letters, and 12 probe letters were distributed along the perimeter of an imaginary clock face that corresponded to the display. Only one of the 12 probe letters was a member of the memory set. The subject's task was to scan around the clock face (starting at 12 o'clock and proceeding clockwise) until he found the positive probe.

The duration of the time spent fixating each negative probe item increased linearly with the memory set size, at a rate of about 50 msec/item. This is compatible with the explanation that each probe item was serially compared to each of the memory set items, and as the memory set size increased, the probe had to be compared to more items in memory. The eye fixated the

probe while the comparison operations occurred. The importance of this finding is that the parameter of 50 msec per item, inferred from the duration of the gaze, is very close to the time of 38 msec per item inferred from reaction time studies with only a single probe (Sternberg, 1969). The duration of fixation on the positive probe item also increased as a function of the memory set size. Of particular importance is that the rate of increase in this duration, the slope, was the same for the positive probe as for the negative probes. This result suggests that the scan through the memory set is exhaustive.

Gould's results provide an important validation for the eye fixation methodology. The eye fixation measures yield results that are completely consistent with Sternberg's careful reaction-time studies. Gould's study, as well as the current research, combines the use of additive factors techniques and eye fixation measures to estimate the time taken to execute a mental operation.

The only eye fixation research that reports a lack of correlation between fixation duration and performance concerns memory for pictures (Loftus, 1972). Loftus found that during learning, the number of fixations, not their total duration, was the best predictor of subsequent recognition memory. However, Tversky (1974) has recently found a positive correlation between the duration of individual eye fixations and later memory. Tversky suggests that the critical variable in relating fixation duration to picture memory might be the kind of features being encoded in the learning phase. The present research does not attempt to account for eye fixations in picture scanning and recognition (for relevant work on this topic, see Buswell, 1935; Mackworth & Bruner, 1970; Mackworth & Morandi, 1967; Noton & Stark, 1971; Potter & Levy, 1969). Our concern has been with ongoing computation

rather than search processes in long-term memory, so the picture recognition issue would take our discussion too far afield. However, the present theoretical framework suggests that the resolution of this problem requires a model of what is encoded during the initial learning and what tests are made during the subsequent recognition phase. As yet, the data relating fixation duration to recognition memory are insufficient to construct a complete model of the mental processes in picture memorization and retrieval.

We have related mental operations to the duration of gaze, not necessarily to the duration of individual fixations. Fixations must be aggregated into a unit consistent with the underlying theory. For example, in the sentence verification task, when a person looked at the plus, whether he made one or two fixations was irrelevant to the fact that he was looking at that piece of information and no other. We measured how long he looked at the plus, not how many fixations actually went into that gaze time. The duration of gaze was systematically related to the duration of underlying processes. The duration of individual fixations may also be so related, but only under those conditions in which each fixation is a psychologically separate unit (cf. also Carpenter & Just, in press; Russo & Rosen, 1975).

Task conditions that optimize the use of eye fixations. The locus of fixation is not always synonymous with the direction of attention. Subjects can be instructed to fixate one referent while attending elsewhere. The possibility of such disassociation makes it important to specify the conditions under which eye fixations are an accurate reflection of what is being processed. One of the most important conditions is that the task require that information from the visual environment be encoded and processed. If the visual display is not relevant, there are no mapping rules between what is being fixated and what is being internally processed. A second condition

is that the task goals be specified for the subject. Asking subjects simply to look at a picture or read some prose permits them to adopt their own definitions of what processing is required and this again makes it difficult to infer the relationship between eye fixations and underlying mental processes. And of course, speeded tasks discourage extraneous processing and the concomitant extraneous fixations.

Some of the rules that govern fixations are general scanning strategies, while other rules are highly specific to the processing in the task being performed. Eye fixations will reveal the mental processes in a particular task only if the task structure minimizes the use of general scanning strategies. An example of this structuring is evident in the study of how people looked at pictures after reading sentences involving affirmative quantifiers e.g., A small proportion of the dots are red, or negative quantifiers, e.g., Few of the dots are red (Carpenter & Just, 1972). The pictures always had a small subset of dots at the top and a large subset at the bottom. Thus, the subject knew to look at the top or at the bottom, depending on whether he wanted to determine the color of the small subset or the large one. This task structure eliminated the need first to search for the desired subset and then to encode its color. The relation between eye fixations and mental operations is even clearer when the role of peripheral information is controlled. The extreme case of this is the computer-based "tunnel vision" in the sentence verification task, in which there is no peripheral information, so the duration of gaze at any locus cannot reflect encoding of information from another locus. These features of the task structure minimize the role of general scanning strategies and thereby make the design more sensitive to the cognitive processes of interest.

In all of these tasks, the eye scan is very much goal directed, in fact, directed by the information present "at the top of the stack." There

are two possible sources of such information, namely, the task environment information computed during the trial. Both sources influenced fixations in the sentence verification task since the instructions to fixate the object determined the first fixation, but the locus of the second fixation was determined by information computed during the trial. After the word in a *The Cat's Hat*) was fixated, the discriminant was in the word determined the locus of the next fixation, in this case, hats. Since both the task structure and the ongoing processing can determine the locus of fixation, both factors must be taken into account in developing a complete processing model.

One domain of eye fixation research that has been neglected by the discipline of task analysis is the area of reading. While there have been many promising empirical studies of eye fixations in reading (cf. Buswell, 1922, 1937; Goldberg, 1975; Gribble, 1975; Levin & Kaplan, 1970; Mackworth, 1976; McConkie & Rayner, 1974; Oakhill, Bower, & Carey, 1987; Tinker, 1958), there is no convergence on a theory of reading. The difficulty is that reading is not a single "reading process", because we read differently in different situations. For example, a newspaper article is read differently than a legal contract, and the same contract is read differently depending on whether one is looking for typographical errors or buying a house. One possible solution to this difficulty is to systematically examine a number of well-defined reading tasks that are of inherent interest. For example, we are currently examining eye fixations of subjects who read and write letters (see Levin, Kaplan, & Grossen, in press, for a description of this task). In another study, we are examining subjects' eye fixations as they read a passage and answer questions about it in a procedural reading comprehension test. In order to develop models of reading, it will probably be necessary to study performance in a number of well-under-

exact task environments, so as to minimize the influence of the stimulus basis on the coding process.

Representation of the problem. Various kinds of questions in the literature appear to involve systematic manipulation of the nature of the coding of the represented problem, for it is always implied, but this aspect is rarely examined in detail. In these cases in which the representations are changed, conclusions should rest on the theoretical assumptions that govern processing of problems in well-defined domains. The conclusion by Hayes-Roth and Hayes-Roth (1979), that in comparing two problems the similarity, necessary for each subject to be subject-free, is determined by an invariant creation.

The evidence from recent psychological research demonstrates that one cannot look solely to information from different sources (like problems or memory tasks) due to processed similarity. Such [illegible] which [illegible] quantitative measures like [illegible] (1979) which suggest to a process that included not just mixed figures. The research process of [illegible] the [illegible] process which [illegible] (e.g. [illegible] Hayes, 1979). In fact, although the relevant information can account from a process to the rule and removed from memory tasks to the other, both sets of data could be explained in terms of the same [illegible] (Hunt, 1978). Obviously, the tasks involving images involve different processes, but to this and various other cases (cf. Kosslyn & Hunt, 1977), the information needs to be [illegible] that similarity must be in part the encoding stages. This suggests that processing aspects of these subsequent stages derived from the [illegible] studies due processes to non-linear features.

Individual [illegible] can rise to [illegible] comparison of the relevant materials. When subjects are deciding whether a visually presented, so-

[illegible] [illegible] [illegible] [

[illegible]

[illegible]

[illegible]

Shepard, R. N., Kilpatric, D. W., & Cunningham, J. P. The internal representation of numbers. Cognitive Psychology, 1975, 7, 82-138.

Shepard, R., & Metzler, J. Mental rotation of three-dimensional objects. Science, 1971, 171, 701-703.

Simon, H. A., & Barenfeld, M. Information processing analysis of perceptual processes in problem solving. Psychological Review, 1969, 76, 473-483.

Sternberg, S. The discovery of processing stages: Extensions of Donder's method. In W. G. Koster (Ed.), Attention and Performance II. Acta Psychologica, 1969, 30, 276-315.

Tichomirov, G. K., & Poznyanskaya, E. D. An investigation of visual search as a means of analyzing heuristics. Voprosy Psikhologii, 1966, 12, 39-53.

Tinker, M. A. Recent studies of eye movements in reading. Psychological Bulletin, 1958, 55, 215-230.

Trabasso, T., Rollins, H., & Shaughnessy, E. Storage and verification stages in processing concepts. Cognitive Psychology, 1971, 2, 239-289.

Tversky, B. Eye fixations in prediction of recognition and recall. Memory & Cognition, 1974, 2, 275-278.

Wickelgren, W. A. Speed-accuracy tradeoff and information processing dynamics. Unpublished manuscript. University of Oregon, Eugene, Oregon, 1974.

Williams, L. G. The effect of target specification on objects fixated during visual search. Perception and Psychophysics, 1966, 1, 315-318.

Winikoff, A. Eye movements as an aid to protocol analysis of problem solving behavior. Unpublished doctoral dissertation, Carnegie-Mellon University, Pittsburgh, Pa., 1967.

Woodworth, R. S. Experimental Psychology. New York: Henry Holt, 1938.

Yarbus, A. L. Eye Movements and Vision. New York: Plenum Press, 1967.

FOOTNOTES

[1]The order of authors is arbitrary. This paper represents a collaborative effort.

We are grateful to Roger Shepard and Herbert Simon for their comments on carlier drafts of this paper. We also thank Roger Shepard for providing copies of the stimulus figures used in the rotation experiment.

The project was supported in part by Research Grant NIE-G-74-0016 from the National Institute of Education, U. S. Department of Health, Education, and Welfare, and Grant MH-07722 from the National Institute of Mental Health.

Requests for reprints should be sent to Marcel Adam Just, Psychology Department, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213.

[2]Metzler and Shepard (1974) report some preliminary observations on the eye movements of two subjects performing the mental rotation task; their subjects also looked back and forth between the two figures.

[3]The 50° steps indicated by our data are suggestively close to 45°, which has more intuitive appeal.

[4]The analysis is based on cell means estimated from a graph of the latencies for comparing random configurations of dots (Gillman & Buckley, 1973). The aggregated data appear in Buckley and Gillman (1974).